Determining the order of differencing in seasonal time series processes

PHILIP HANS FRANSES[†] AND A. M. ROBERT TAYLOR[‡]

[†]*Econometric Institute, Erasmus University Rotterdam, The Netherlands*
E-mail: `franses@few.eur.nl`
[‡]*Department of Economics, University of Birmingham, Birmingham, UK*
E-mail: `R.Taylor@bham.ac.uk`

Summary   In this paper we propose a sequential testing procedure to determine the order of differencing in seasonally observed time series processes, which builds upon existing approaches developed for nonseasonal series. We allow for the possible presence of multiple unit roots at both the zero and seasonal frequencies of the data. Multiple seasonal unit roots can be useful in circumstances where one does not wish to take logarithms of the data-set at hand. Multiple unit roots at the seasonal frequencies also appear in commonly applied seasonal adjustment filters such as Census X-11. The testing procedure developed in this paper can therefore be used to obtain some *a priori* insight into the likely properties of seasonally adjusted time series. In particular it may be used to investigate whether or not the seasonally adjusted series can be expected to be strictly noninvertible at the seasonal frequencies of the data. The proposed testing procedure is shown be asymptotically consistent. The unit root statistics arising at each stage of the testing procedure are shown to have familiar limiting null distributions so that, at least to an approximation, existing critical values may be used. Empirical applications are provided to illustrate the practical usefulness of our approach.

Keywords: *Multiple seasonal unit roots, Seasonal adjustment.*

1. INTRODUCTION

In this paper we consider the problem of testing for multiple unit roots in univariate seasonal time series processes. We allow for such roots at both the zero and seasonal frequencies. Hylleberg *et al.* (1990) (hereafter referred to as HEGY) develop regression-based tests for seasonal unit roots in a quarterly context as separate t and F-tests for unit roots at each of the zero and seasonal frequencies against stationary alternatives. They consider regressions which may include an intercept, seasonal intercepts and a trend variable, while Smith and Taylor (1998) (ST1) allow for seasonally varying trends. Monthly versions of the HEGY procedure have been developed by Beaulieu (1993) (BM) and Franses (1991), while Smith and Taylor (1999a) (ST2) develop HEGY-type tests for an arbitrary seasonal aspect. Other statistics relevant to these procedures have also been developed in, among others, Ghysels *et al.* (1994) (GLN) and Taylor (1998).

We extend the HEGY framework in order to enable the practitioner to test for multiple unit roots at both the zero and seasonal frequencies. We do so in exactly the same manner as Dickey and Pantula (1987) (DP) and Pantula (1989) extend the standard Dickey and Fuller (1979) (DF)

(non-seasonal) unit root test to determine the appropriate order of differencing in processes with a potentially higher order of integration than one. Haldrup (1994) gives formal asymptotic results which demonstrate the need to take care of such higher orders. Attention in the paper often focuses on the quarterly and monthly settings, and on the possibility of double unit roots at each of the zero and seasonal frequencies. The procedure which we detail is, however, sufficiently general to handle data with other periodicities, e.g. bi-monthly or bi-annual data, and with potentially higher orders of integration.

The outline of our paper is as follows. Section 2 outlines the motivation for this paper in some detail. Section 3 gives the sequential multiple testing approach, presenting a general algorithm. Representations for the limiting distributions of these statistics are also provided in Section 3. Section 4 provides applications of the sequential testing procedure to quarterly M4 in the UK and monthly industrial production in the US. Section 5 concludes the paper.

2. MOTIVATION

Consider the finite-order autoregressive (AR) time series process $\{y_t\}_{t=1}^T$, observed with constant periodicity S, which admits the DGP:

$$\phi_p(L)(y_t - \mu_t^*) = \epsilon_t \iff \phi_p(L)y_t = \mu_t + \epsilon_t \qquad (2.1)$$

where μ_t^* and μ_t are functions of t, linear in a vector of parameters, such that $\mu_t = \phi_p(L)\mu_t^*$, $\phi_p(L) \equiv 1 - \sum_{j=1}^p \phi_j L^j$ an AR polynomial of order p in the lag operator L, $L^k y_t = y_{t-k}$, $k = 0, 1, \ldots$. We postpone an exact parametrization of μ_t until Section 3. The disturbance process $\{\epsilon_t\}$ is assumed to be a martingale difference sequence satisfying the conditions in Chan and Wei (1998, pp. 367–368). The initial conditions $\{y_j\}_{j=1-p}^0$ are assumed fixed. Notice that the mean of y_t is independent of the form of $\phi_p(L)$, subject to certain restrictions on the initial conditions. In order to simplify notation we impose S to be even. Generalizations to allow for odd S are immediate but require a somewhat more involved notation; see ST2.

In this paper we consider the factorization of $\phi_p(L)$ of DGP (2.1) into,

$$\phi_p(L) = \alpha(L)\beta(L) \qquad (2.2)$$

where $\alpha(L) = 1 - \sum_{i=1}^{p*} \alpha_i L^i$, $0 \leq p* \leq p$, is such that all the roots of $\alpha(z^{-1}) = 0$ lie on the unit circle at some or all of the frequencies $\omega_k = 2\pi k/S$, $k = 0, \ldots, S - 1$. In contrast, $\beta(L) = 1 - \sum_{k=1}^{p-p*} \beta_k L^k$ is such that all the roots of $\beta(z^{-1}) = 0$ lie inside the unit circle; that is, while $\beta(L)$ is a stationary polynomial, $\alpha(L)$ contains all roots of length unity, occurring at some subset of the zero and seasonal frequencies. To ensure that $\{y_t\}$ is a real-valued series, both $\alpha(L)$ and $\beta(L)$ must be real. Consequently, roots at the harmonic frequencies, $\omega = 2\pi k/S$, $k = 1, \ldots, (S/2) - 1$, occur as conjugate pairs; cf. HEGY. We consider the factorization of the familiar seasonal difference operator $(1 - L^S)$ into $\prod_{k=0}^{S-1} (1 - e^{i2\pi k/S}L)$, to yield the *flexible filter* representation for $\alpha(L)$:

$$\alpha(L) = \prod_{k=0}^{S/2} \alpha_k(L) \qquad (2.3)$$

where $\alpha_0(L) \equiv (1 - L)^{d_0}$, $\alpha_k(L) \equiv \{1 - 2\cos(2\pi k/S)L + L^2\}^{d_k}$, $k = 1, \ldots, (S/2) - 1$ and $\alpha_{S/2}(L) \equiv (1 + L)^{d_{S/2}}$. In each case $d_j \in \{0, 1, 2, \ldots\}$, $j = 0, \ldots, S/2$. The process $\{y_t\}$ of

(2.1) is then said to be integrated of order d_j at frequency $\omega_j = 2\pi j/S$, $j = 0, \ldots, S/2$, denoted $x_t \sim I_j(d_j)$; cf. HEGY.

Under the assumption that $d_j \in \{0, 1\}$, $j = 0, \ldots, S/2$, and in order to investigate for the presence or otherwise of zero- and seasonal-frequency unit roots in $\{y_t\}$ in the case of quarterly data ($S = 4$), HEGY take a first-order (local) expansion of $\phi_p(L)$ about the (unit) roots ± 1, $\pm i$. Substitution of the local expansion into (2.1) yields the test (auxiliary) regression,

$$\phi_{p-4}(L)\Delta_4 y_t = \mu_t + \sum_{j=1}^{4} \pi_j y_{j,t-1} + \epsilon_t, \tag{2.4}$$

an exact linear re-parametrization of (2.1); cf. HEGY. In the context of (2.4), $\phi_{p-4}(L)$ is a stationary AR lag (augmentation) polynomial of order $p - 4$, $\Delta_S^k \equiv (1 - L^S)^k$, and

$$y_{1,t} \equiv (1 + L + L^2 + L^3)y_t, \quad y_{2,t} \equiv -(1 - L + L^2 - L^3)y_t,$$
$$y_{3,t} \equiv -L(1 - L^2)y_t, \quad y_{4,t} \equiv -(1 - L^2)y_t. \tag{2.5}$$

It is the $\{\pi_i\}_{i=1}^4$ parametrs of (2.4) which are of focal interest. As demonstrated in HEGY, a zero-frequency unit (stable) root is obtained under $\pi_1 = 0$ ($\pi_1 < 0$); $\pi_2 = 0$ ($\pi_2 < 0$) yields a bi-annual (Nyquist) frequency unit (stable) root, while a complex conjugate pair of unit (stable or no) roots at the annual (seasonal harmonic) frequencies implies $\pi_3 = \pi_4 = 0$ ($\pi_3 < 0 \cup \pi_4 \neq 0$). In the case where (2.4) contains no deterministic regressors, Chan and Wei (1998) demonstrate that the limiting null distribution of the regression t-statistics for $\pi_1 = 0$ and $\pi_2 = 0$ (henceforth t_1 and t_2, respectively) from (2.4), is that of the standard Dickey–Fuller distribution (see Fuller (1996, Table 10.A.2, p. 642)). Engle *et al.* (1993) give the asymptotic null distributions of the regression t-statistics for $\pi_4 = 0$ (t_4) and, conditional on $\pi_4 = 0$, $\pi_3 = 0$ (t_3), and the joint F-test for $\pi_3 = \pi_4 = 0$ (F_{34}). GLN derive asymptotic representations for additional joint F-tests. A complete typology of distributional results pertinent to (2.4) appears in ST1, with generalizations to arbitrary S in ST2.

Given the asymptotic and Monte Carlo results in Sen (1985), Pantula (1989) and Haldrup (1994) for the non-seasonal DF test, the aforementioned asymptotic results can only be expected to be valid when the assumption that $d_j \in \{0, 1\}$, $j = 0, \ldots, S/2$, holds. In other words, when $\{y_t\}$ has multiple unit roots, the commonly applied asymptotic theory will not hold, and one may expect inappropriate inference. To illustrate the effects of basing seasonal unit root test-statistics on (2.4), setting μ_t to contain seasonal dummies and a linear trend, when the process has multiple unit roots, we used Monte Carlo simulation to generate 40 000 series, each of length 200, according to the DGP: $\Delta_4^j y_t = \epsilon_t \sim IN(0, 1)$, $j = 1, 2$. All experiments were programmed using the RNDN random number generator of GAUSS 3.1. When $j = 1$, the t_1, t_2 and F_{34} HEGY test-statistics all rejected the relevant unit root null at very close to 5% of the time when using the relevant 5% critical values from HEGY. In contrast, for $j = 2$, the t_1 test rejected the zero-frequency unit root null (in favour of stationarity) 14% of the time, the t_2 test rejected the bi-annual frequency unit root 7% of the time, and the F_{34} test rejected the annual frequency unit root 88% of the time, even though the process was integrated of order two at each of the zero, bi-annual and annual frequencies. These rejection frequencies rose to 21, 9 and 94% respectively, for the t_1, t_2 and F_{34} tests when the sample size was increased to 400. Intuitively, we would hope that the unit root null would be rejected no more than 5% of the time (in favour of stationarity) when the process contained two or more unit roots at the frequency under investigation. Our simulation evidence, however, shows the converse to be true. Such results may lead to a deterioration of

out-of-sample forecasting performance. The empirical results in Clements and Hendry (1997) for quarterly time series, and Franses (1991) for monthly time series show that the choice of differencing filter can be crucial for out-of-sample forecasts.

An alternative testing strategy, which allows for two zero-frequency unit roots in $\{y_t\}$, has been proposed by Osborn *et al.* (1988) (henceforth OCSB). This approach builds upon the work of Hasza and Fuller (1982). Consider again (2.1) with $\phi_p(L) = (1 - \lambda_1 L)(1 - \lambda_S L^S)$. OCSB show that a Taylor series expansion of $\phi_p(L)$ about λ_1 and λ_S yields the test regression,

$$\Delta_1 \Delta_S y_t = \mu_t + \delta_1 \Delta_S y_{t-1} + \delta_S \Delta_1 y_{t-S} + \epsilon_t, \tag{2.6}$$

such that $\delta_1 \equiv \lambda_1 - 1$, $\delta_S \equiv \lambda_S - 1$, with $\delta_1 \delta_S = 0$ ($\lambda_1 = 1$ and/or $\lambda_S = 1$) maintained. Where $\lambda_1 = 1$ ($\delta_1 = 0$), $\{\Delta_1 y_t\}$ is a stable AR(S) process if $-1 < \lambda_S < 1$ ($-2 < \delta_S < 0$). Conversely, if $\lambda_S = 1$ ($\delta_S = 0$) and $-1 < \lambda_1 < 1$ ($-2 < \delta_1 < 0$), then $\{\Delta_S y_t\}$ is a stable AR(1) process. Finally, if $\lambda_1 = \lambda_S = 1$, then the double differencing filter $\Delta_1 \Delta_S$ is required to transform $\{y_t\}$ to a stochastically stationary process. It is clear that the OCSB approach suffers from two obvious drawbacks. Firstly, the framework adopts a maintained hypothesis that no more than one of λ_1 and λ_S may be less than unity, thereby imposing the condition that $y_t \sim I_0(d)$, $d \in \{1, 2\}$, since $\Delta_S \equiv \Delta_1(1 + L + \cdots + L^{S-1})$. Secondly, by focusing on a single testable parameter, λ_S, the procedure imposes the condition that the modulus of roots at the seasonal frequencies $\omega_j = 2\pi j/S$, $j = 1, \ldots, S-1$, are equal. Both assumptions are violated, for example, by the quarterly ($S = 4$) process, $(1 + L) y_t = \epsilon_t$, which is simultaneously $I_0(0)$, $I_1(0)$ and $I_2(1)$. Where either, or both, of these maintained hypotheses are violated the OCSB testing strategy will be inefficient, relative to a procedure which does not impose such restrictions; cf. Section 4.1.

In addition to statistical motivations for a testing approach allowing for multiple unit roots at both the zero and seasonal frequencies, there are also two practical motivations. The first concerns the possible presence of increasing seasonal variation in time series. Nowadays it is current practice to take natural logarithmic transforms of the $\{y_t\}$ series in order to establish approximately constant seasonal variation. Although seasonal trends, such as in (2.4), can also accommodate increasing seasonal variation, such a parametrization is not often considered. An alternative model that can generate increasing seasonality is given by allowing for double filters such as Δ_4^2 for quarterly data, see, e.g. Bowerman *et al.* (1990). Hence, if one does not wish to take logs, one may entertain model specifications which assume multiple unit roots at the seasonal frequencies. It should be stressed, however, that the foregoing models are strictly non-nested, reinforcing the usual requirement on the practitioner to test the statistical assumptions which underpin the regression models from which the unit root tests are derived; e.g. linearity, homoscedasticity, serial uncorrelatedness. For example, as pointed out by a referee, if $\Delta_S y_t = \epsilon_t \sim NID(0, \sigma^2)$, then $\Delta_S^2 Y_t$, where $Y_t = \exp(y_t)$, will neither be a linear process, nor will it admit a stationary and invertible $ARMA$ representation; cf. (2.1).

Finally, an important motivation for testing for multiple seasonal unit roots is that it provides a method to evaluate the possible properties of seasonally adjusted data. Many seasonal adjustment methods, involve applying a linear symmetric moving average filter $W(L)$ to the data that includes the factor $S(L) \equiv (1 + L + L^2 + \cdots + L^{S-1})$, or even $S(L)^2$, which implicitly assumes the presence of (double) seasonal unit roots. For example, Laroque (1977, Table 3) demonstrates that in the case of the linear Census X-11 method, $W(L)$ contains the factor $S(L)^2$. When a time series does not admit these unit roots, the resulting adjusted series will be non-invertible; see Maravall (1995) for an extensive discussion. In order to obtain some *a priori* insight in the properties of seasonally adjusted time series, one may use our test procedure, which we present in the next section, to check whether or not the $S(L)$, or indeed $S(L)^2$, component is present.

3. A SEQUENTIAL TESTING PROCEDURE FOR MULTIPLE SEASONAL UNIT ROOTS

Consider again (2.1) with μ_t taken to be

$$\mu_t = \sum_{q=1}^{S} \gamma_q^* D_{q,t} - \sum_{q=1}^{S} \delta_q^* D_{q,t} t, \tag{3.1}$$

$$= \sum_{k=0}^{S/2} (\gamma_k^{*\alpha} \sin t\omega_k + \gamma_k^{*\beta} \cos t\omega_k) - \sum_{k=0}^{S/2} (\delta_k^{*\alpha} \sin t\omega_k + \delta_k^{*\beta} \cos t\omega_k) t, \tag{3.2}$$

where the $D_{q,t}$, $q = 1, \ldots, S$, are a set of conventional seasonal indicator variables. The specification (2.1), (3.1) therefore allows the deterministic mean of y_t to be a polynomial in t of order one whose intercept and slope coefficients may differ across seasons *via* the parameters γ_q^* and δ_q^* respectively, $q = 1, \ldots, S$; cf. ST1 and Taylor (1998) for $S = 4$ and $S = 12$, respectively. We may identify other specifications for the mean function of interest within (3.1) which are special cases thereof: (i) zero intercept, zero slope: $\gamma_q^* = \delta_q^* = 0$, $q = 1, \ldots, S$; (ii) fixed intercept, zero slope: $\gamma_q^* = \gamma$, $\delta_q^* = 0$, $q = 1, \ldots, S$; (iii) seasonally varying intercept, zero slope: $\delta_q^* = 0$, $q = 1, \ldots, S$; (iv) fixed intercept, fixed slope: $\gamma_q^* = \gamma$, $\delta_q^* = \delta$, $q = 1, \ldots, S$; (v) seasonally varying intercept, fixed slope $\delta_q^* = \delta$, $q = 1, \ldots, S$; cf. HEGY and BM for $S = 4$ and $S = 12$, respectively. The re-parametrization of (3.1) embodied in (3.2) associates the parameters $\gamma_k^{*\alpha}, \gamma_k^{*\beta}$ and $\delta_k^{*\alpha}, \delta_k^{*\beta}$ with the corresponding seasonal frequency $\omega_k \equiv 2\pi k/S$, $k = 0, \ldots, S/2$. Consequently, we term $\gamma_k^{*\alpha}, \gamma_k^{*\beta}$ and $\delta_k^{*\alpha}, \delta_k^{*\beta}$ *spectral intercept* and *spectral time-trend* parameters respectively. See, for example, Hannan *et al.* (1970) for a similar re-parametrization.

We may re-express the AR(p) process (2.1), (3.2) alternatively as:

$$\phi_p(L)y_t = \sum_{k=0}^{S/2} (\gamma_k^{\alpha} \sin t\omega_k + \gamma_k^{\beta} \cos t\omega_k) + \sum_{k=0}^{S/2} (\delta_k^{\alpha} \sin t\omega_k + \delta_k^{\beta} \cos t\omega_k) t + \epsilon_t, \tag{3.3}$$

where the $\{\gamma_k^{\alpha}, \gamma_k^{\beta}, \delta_k^{\alpha}, \delta_k^{\beta}\}_{k=0}^{S/2}$ parameters of (3.3) are functions of the $\{\gamma_k^{*\alpha}, \gamma_k^{*\beta}, \delta_k^{*\alpha}, \delta_k^{*\beta}\}_{k=0}^{S/2}$ and $\{\phi_j\}_{j=1}^{p}$ parameters of (2.1), (3.2); see Smith and Taylor (1999b, p. 456) for the exact relationships. Crucially, $\delta_k^{\alpha} = \delta_k^{\beta} = 0$, if $\alpha(L)$ of (2.3) contains the factor $(1 - e^{i2\pi k/S}L)$, or indeed if the data is transformed by some filter containing this factor, while $\gamma_k^{\alpha} = \gamma_k^{\beta} = 0$ additionally if $\alpha(L)$ admits the factor $(1 - e^{i2\pi k/S}L)^j$, $j = 2, 3, \ldots, k \in \{0, \ldots, S/2\}$. Moreover, special cases (i) to (iv) of (3.1) also impose the following restrictions on (3.3): (i) $\gamma_k^{\alpha} = \gamma_k^{\beta} = 0$, $\delta_k^{\alpha} = \delta_k^{\beta} = 0, k = 0, \ldots, S/2$; (ii) $\gamma_k^{\alpha} = \gamma_k^{\beta} = 0, k = 1, \ldots, S/2, \delta_k^{\alpha} = \delta_k^{\beta} = 0, k = 0, \ldots, S/2$; (iii) $\delta_k^{\alpha} = \delta_k^{\beta} = 0$, $k = 0, \ldots, S/2$; (iv) $\gamma_k^{\alpha} = \gamma_k^{\beta} = 0$, $k = 1, \ldots, S/2$, $\delta_k^{\alpha} = \delta_k^{\beta} = 0$, $k = 1, \ldots, S/2$; (v) $\delta_k^{\alpha} = \delta_k^{\beta} = 0, k = 1, \ldots, S/2$. In what follows, we shall use μ_t as a generic term, indicating the set of deterministic variables appearing with non-zero coefficients in (3.3), and in our subsequent test regressions. The inclusion of these variables renders the resulting unit root tests similar (both exactly and asymptotically) with respect to the nuisance parameters $\{\gamma_k^{*\alpha}, \gamma_k^{*\beta}, \delta_k^{*\alpha}, \delta_k^{*\beta}\}_{k=0}^{[S/2]}$ of (3.2), or equally $\{\gamma_q^*, \delta_q^*\}_{q=1}^{S}$ of (2.1). As we have just seen, the exact form of μ_t will therefore depend both on the specification of the mean function of y_t, and on the specific null and alternative hypotheses under test. It is advisable, however, that μ_t contain, as a

minimum, spectral intercepts in order that the tests are (exact) similar to the initial values of the process.

As discussed in Section 2, published work to date has focused exclusively on the choice between $d_j = 1$ and $d_j = 0$, $j = 0, \ldots, S/2$; that is, between processes that are respectively once seasonally integrated and stochastically stationary at frequency $\omega_j = 2\pi j/S$, $j = 0, \ldots, S/2$. In this paper we widen the choice to that between $d_j = 0, 1, \ldots k$, k being some notional upper-bound on the maximum order of integration allowed. Initially, purely to ease notation, we shall assume that k is set common to all frequencies. We then provide a simple example for $S = 4$ which relaxes this condition. In practice, we would be unlikely to consider any $k > 2$. For $k = 2$ we may both investigate the issues relating to the seasonal adjustment of time series as discussed in Section 2, and provide an alternative to the OCSB approach.

In order to determine the appropriate order of differencing required at both the zero and seasonal frequencies of the time series at hand, we adopt a sequential testing procedure, developed along parallel lines to the non-seasonal procedure of DP and Pantula (1989). The testing procedure developed by those authors is specifically designed so that at each step of the sequential procedure one tests a hypothesis structure which parallels that of the conventional DF test, maintaining the asymptotic null distribution of that test-statistic. Specifically, they consider the sequence of regressions

$$\Delta_1^k y_t = \sum_{i=1}^{\tau} \rho_{k+1-\tau} \Delta_1^{k-i} y_{t-1} + \sum_{j=1}^{p-k} \beta_{j,k+1-\tau} \Delta_1^k y_{t-j} + \mu_t + \epsilon_{t,\tau}, \qquad (3.4)$$

for $\tau = 1, \ldots, k$, k being the notional upper bound on the order of integration at the zero frequency; see Pantula (1989, p. 260, Equation (2.7)). The first stage of the procedure, $\tau = 1$, tests the null hypothesis that $y_t \sim I_0(k)$ against the alternative that $y_t \sim I_0(k-1)$, under the maintained hypothesis that $y_t \sim I_0(d)$, $d \geq k - 1$, *via* the usual DF-type test on ρ_k. Under the null, this statistic has the usual DF limiting distribution (for the given form of μ_t), and is consistent if $y_t \sim I_0(b)$, $b < k$. If one rejects the null, one progresses to the second stage of the procedure, $\tau = 2$ in (3.4). Here we test the null hypothesis that $y_t \sim I_0(k-1)$ against the alternative that $y_t \sim I_0(k-2)$, the maintained hypothesis now that $y_t \sim I_0(d)$, $d \geq k - 2$, again *via* the DF test on ρ_{k-1}. The procedure continues until we cannot reject the null at any particular stage of the procedure, or until $\tau = k$. At any given stage of the procedure, if the maintained hypothesis holds, then (3.4) constitutes a linear re-parametrization of (3.3) with $\epsilon_{t,\tau} \equiv \epsilon_t$, for all t. Although violation of the maintained hypothesis will impose invalid restrictions on the estimated model (in the unrestricted case the first summation term in (3.4) would run from 1 to k, rather than 1 to τ, for all τ), the associated error process $\{\epsilon_{t,\tau}\}$ being non-invertible at frequency zero, DP demonstrate that using the sequence of regressions in (3.4) yields considerably more powerful tests.

We now detail our sequential testing procedure, generalizing the non-seasonal procedure of DP and Pantula (1989), outlined above, to allow for multiple unit roots at both the zero and seasonal frequencies. Consequently, at least to an approximation, no new tables of finite-sample critical values will be required. Asymptotic critical values will be unchanged.

Step 1a:.   Regress $\Delta_S^k y_t$ on $\{\Delta_S^k y_{t-j}\}_{j=1}^{p-Sk}$, $\{\Delta_S^{k-1} y_{j,t-1}\}_{j=1}^{S}$, and a set of deterministic variables in order to obtain the OLS estimates of the focal parameters $\{\pi_{j,k}\}_{j=1}^{S}$ from the first-stage auxiliary

regression,

$$\phi_{p-Sk,k}(L)\Delta_S^k y_t = \sum_{j=1}^{S} \pi_{j,k}\Delta_S^{k-1} y_{j,t-1} + \mu_t + \epsilon_{t,1}, \tag{3.5}$$

where $\phi_{p-Sk,k}(L)$ is a stationary polynomial of order $p - Sk$. The $\{y_{j,t}\}_{j=1}^{S}$ of (3.5) constitute a set of S linear transformations of y_t which separate out the zero and seasonal frequencies of the data; see ST2 for a general formulation, and for the special cases of $S = 4$ and $S = 12$, see (2.5) of this paper and pp. 307–308 of BM, respectively.

Remark 1. *In the case of $k = 1$, (3.5) reduces to the usual auxiliary, or HEGY, regression for testing the conventional seasonal unit root hypothesis; e.g. for $S = 4$ see (2.4) of this paper. More generally, (3.5) may be viewed as the HEGY regression applied to the series $\{\Delta_S^{k-1} y_t\}$.*

Step 1b: Following HEGY, compute the S regression-based t-statistics $t_{j,k}$ for,

$$H_{0,j}^k : \pi_{j,k} = 0 \qquad j = 1, \ldots, S, \tag{3.6}$$

against

$$H_{1,j}^k : \pi_{j,k} < 0 \qquad j = 1, 2, \tag{3.7}$$

$$H_{1,2j+1}^k : \pi_{2j+1,k} < 0 \qquad j = 1, \ldots, S/2 - 1, \tag{3.8}$$

$$H_{1,2j+2}^k : \pi_{2j+2,k} \neq 0 \qquad j = 1, \ldots, S/2 - 1. \tag{3.9}$$

For each test-statistic, our decision rule is exactly as in the conventional seasonal unit root case. That is, we reject $H_{0,h}^k : \pi_{h,k} = 0$, $h = 1, 2$, if $t_{h,k} < t_h^*(\alpha, T)$, while we reject $\cap_{i=1}^2 H_{0,2j+i}^k$ if either $t_{2j+2,k} < t_{2j+2}^*(\alpha/2, T)$ or $t_{2j+2,k} > t_{2j+2}^*(1 - \alpha/2, T)$, or, conditional on the acceptance of $H_{0,2j+2}^k, t_{2j+1,k} < t_{2j+1}^*(\alpha, T)$, $j = 1, \ldots, S/2 - 1$; cf. HEGY. In each case, $t_j^*(\alpha, T)$ denotes the lower α percentile of the conventional t_j, $j = 1, \ldots, S$, distribution for a sample of size T taken from the auxiliary regression (3.5) with $k = 1$; see HEGY and ST1 for $S = 4$, and BM and Taylor (1998) for $S = 12$.

Remark 2. *As in HEGY, the hypotheses relating to the seasonal harmonic frequencies, $\omega_j = 2\pi j/S$, $j = 1, \ldots, S/2 - 1$, can also be tested via the F-statistic $F_{j,k}$ for the exclusion of $(\Delta_S^{k-1} y_{2j+1,t-1}, \Delta_S^{k-1} y_{2j+2,t-1})$, $j = 1, \ldots, S/2 - 1$, from (3.5). We may also consider joint-frequency F-statistics for the nullity of some subset of the $\{\pi_{j,k}\}_{j=1}^{S}$ parameters pertinent to more than one frequency of the data. In particular, we consider the F-statistics, $F_{1\ldots S,k}$ and $F_{2\ldots S,k}$ of GLN and ST1 ($S = 4$) and Taylor (1998) ($S = 12$) for the exclusion of the regressors $\{\Delta_S^{k-1} y_{j,t-1}\}_{j=1}^{S}$ and $\{\Delta_S^{k-1} y_{j,t-1}\}_{j=2}^{S}$ respectively, from (3.5).*

Step 1c: In each case we define the indicator variable $\delta_{j,k}$ such that $\delta_{j,k} = 1$ if we do not reject $H_{0,j}^k$, $j = 1, \ldots, S$, while $\delta_{j,k} = 0$ if we reject $H_{0,j}^k$ for $H_{1,j}^k$, $j = 1, 2$, and $\delta_{2j+i,k} = 0$, $i = 1, 2$, if we reject $\cap_{i=1}^2 H_{0,2j+i}^k$, $j = 1, \ldots, S/2 - 1$. Finally we define the set of S *reverse* indicators: $\delta_{i,k}' = 1 - \delta_{i,k}, i = 1, \ldots, S$.

Under $H_{0,1}^k$, $y_t \sim I_0(k)$, while under $H_{1,1}^k$, $y_t \sim I_0(k - 1)$, under the maintained hypothesis that $y_t \sim I_0(b)$, $b \geq k - 1$. Similarly, under $H_{i,2}^k$, $y_t \sim I_{S/2}(k - i)$, $i = 0, 1$, again under the

maintained hypothesis that $y_t \sim I_{S/2}(b)$, $b \geq k - 1$. At the harmonic frequencies, $\omega_j = 2\pi j/S$, $j = 1, \ldots S/2 - 1$, $y_t \sim I_j(k)$ under $\cap_{i=1}^2 H_{0,2j+i}^k$, while $y_t \sim I_j(k-1)$ under $\cup_{i=1}^2 H_{1,2j+i}^k$, with $y_t \sim I_j(b)$, $b \geq k - 1$, maintained. Generalizing the second stage of the DP sequence, we therefore progress in Step 2 to investigate, for frequencies ω_j, $j \in \{0, \ldots, S/2\}$ at which the null hypothesis of k unit roots is rejected, whether or not the data can reject the null hypothesis that $y_t \sim I_j(k-1)$ in favour of the alternative that $y_t \sim I_j(k-2)$.

Step 2a: Regress $\Delta_S^k y_t$ on the regressors $\{\Delta_S^k y_{t-j}\}_{j=1}^{p-Sk}$, $\{\delta_{i,k}' \Delta_S^{k-1} y_{i,t-1}\}_{i=1}^S$ from (3.5) together with the additional regressors, $\{\delta_{i,k}' \Delta_S^{k-2} \delta_k(L) y_{i,t-1}\}_{i=1}^S$, where

$$\delta_k(L) \equiv (1 - L)^{\delta_{1,k}} (1 + L)^{\delta_{2,k}} \times \prod_{j=1}^{S/2-1} \left\{1 - 2\cos(2\pi j/S)L + L^2\right\}^{\delta_{2j+1,k}} \tag{3.10}$$

to obtain the OLS estimates of the focal parameters $\{\pi_{j,k-1}\}$ from the second-stage auxiliary regression,*

$$\phi_{p-Sk,k-1}(L)\Delta_S^k y_t = \sum_{j=1}^S \pi_{j,k} \left(\delta_{j,k}' \Delta_S^{k-1} y_{j,t-1}\right)$$
$$+ \sum_{j=1}^S \pi_{j,k-1} \left\{\delta_{j,k}' \Delta_S^{k-2} \delta_k(L) y_{j,t-1}\right\} + \mu_t + \epsilon_{t,2}. \tag{3.12}$$

Step 2b: Compute, where relevant, the regression t-statistics $t_{j,k-1}$ and F-statistics $F_{j,k-1}$, $j = 1, \ldots, S/2 - 1$, $F_{1 \ldots S, k-1}$ and $F_{2 \ldots S, k-1}$, or any other relevant joint-frequency F-tests, from (3.12), using the same hypothesis structure as in (3.6), (3.9), but replacing k by $k-1$ and adopting the same decision rules as in Step 1b. This implies, taking the zero frequency as an example, that under $H_{0,1}^{k-1}$, $y_t \sim I_0(k-1)$, while under the corresponding alternative hypothesis $H_{1,1}^{k-1}$, $y_t \sim I_0(k-2)$, the maintained hypothesis now being that $y_t \sim I_0(b)$, $b \geq k - 2$.

Step 2c: Define a new set of indicator variables $\{\delta_{j,k-1}\}_{j=1}^S$ and reverse indicator variables $\{\delta_{j,k-1}'\}_{j=1}^S$, in an identical manner to that described in Step 1, with the additional condition that if $\delta_{j,k} = 1$, then so we set $\delta_{j,k-1} = 1$ and $\delta_{j,k-1}' = 0$.

Step 2 is then repeated a maximum of a further $(k-2)$ times. So in Step 3 we would next investigate whether or not the data reject the null hypotheses that $y_t \sim I_j(k-2)$, again for relevant $j \in$

---

*Notice that regression (3.12) may be re-parametrized as

$$\phi_{p-S(k-1)-\lambda,k-1}(L)\delta_k(L)\Delta_S^{k-1} y_t = \sum_{j=1}^S \pi_{j,k-1} \left\{\delta_{j,k}' \Delta_S^{k-2} \delta_k(L) y_{j,t-1}\right\} + \mu_t + \epsilon_{t,2} \tag{3.11}$$

where λ is the order of $\delta_k(L)$. The two regressions given by (3.12) and (3.11) are formally equivalent, differing only in the manner in which they parameterize the lag augmentation variables. In particular, the $\{\pi_{j,k-1}\}$ parameters are identical functions of the $\{\phi_j\}_{j=1}^p$ parameters of (2.1) under both representations. A similar bijection applies in all subsequent steps of the procedure.

$\{0, \ldots, S/2\}$, against the alternative that $y_t \sim I_j(k-3)$. As an example, if a third stage were required, the relevant auxiliary regression would be obtained by regressing $\Delta_S^k y_t$ on $\{\Delta_S^k y_{t-j}\}_{j=1}^{p-Sk}$, $\{\delta'_{i,k}\Delta_S^{k-1} y_{i,t-1}\}_{i=1}^S$, $\{\delta'_{i,k-1}\Delta_S^{k-2}\delta_k(L)y_{i,t-1}\}_{i=1}^S$, and $\{\delta'_{i,k-1}\Delta_S^{k-3}\delta_{k-1}(L)y_{i,t-1}\}_{i=1}^S$ where

$$\delta_{k-1}(L) \equiv \delta_k(L) \times (1-L)^{\delta_{1,k-1}}(1+L)^{\delta_{2,k-1}} \times \prod_{j=1}^{S/2-1} \left\{ 1 - 2\cos(2\pi j/S)L + L^2 \right\}^{\delta_{2j+1,k-1}},$$

(3.13)

with the $\{\delta_{j,k-1}\}_{j=1}^S$ and $\{\delta'_{j,k-1}\}_{j=1}^S$ as defined above. Notice that the procedure effectively terminates if at any step i, $i \in \{1, \ldots, k\}$, in the procedure $\delta_{j,i} = 1$, for all $j = 1, \ldots, S$.

At any given stage of the procedure, if the relevant maintained hypotheses hold, then the given auxiliary regression will constitute a bijective re-parametrization of (3.3) under both the null and alternative hypotheses of interest at that stage (Chan and Wei1998).

Example 1 (quarterly data): Suppose, faced with data generated from the DGP

$$\Delta_4 \left(y_t - \sum_{q=1}^4 \gamma_q^* D_{q,t} - \delta t \right) = u_t, \quad \phi_{p-4}(L)u_t = \epsilon_t \sim IN(0, \sigma^2),$$

$\phi_{p-4}(L)$ a stationary polynomial of order $p-4$, an investigator felt that the process might be integrated of order two at the zero frequency. The testing procedure outlined above may be easily generalised to handle this scenario. To illustrate, we therefore set the upper bound to be two at the zero frequency and one at the seasonal frequencies. The testing procedure would then be as follows.

First estimate by OLS the focal parameters $\pi_{1,2}$ and $\{\pi_{j,1}\}_{j=2}^4$ from the first stage auxiliary regression

$$\Delta_1 \Delta_4 y_t = \pi_{1,2}\Delta_1 y_{1,t-1} + \sum_{j=2}^4 \pi_{j,1}\Delta_1 y_{j,t-1} + \sum_{j=1}^{p-5} \phi_{j,2}\Delta_1\Delta_4 y_{t-j} + \mu_t + \epsilon_t, \qquad (3.14)$$

with μ_t comprising spectral intercepts. Notice that (3.14) is just a HEGY regression for $\{\Delta_1 y_t\}$. In theory we would reject $\mathsf{H}_{0,1}^2 : \pi_{1,2} = 0$, but not reject $\mathsf{H}_{0,j}^1 : \pi_{j,1} = 0$, $j = 2, \ldots, 4$. Consequently, we would set $\delta_{1,2} = 0$, $\delta'_{1,2} = 1$, and $\delta_{j,1} = 1$, $\delta'_{j,1} = 0$, $j = 2, \ldots, 4$. The second stage auxiliary regression would then be of the form:

$$\Delta_1 \Delta_4 y_t = \pi_{1,2}\Delta_1 y_{1,t-1} + \pi_{1,1} y_{1,t-1} + \sum_{j=1}^{p-5} \phi_{j,1}\Delta_1\Delta_4 y_{t-j} + \mu_t + \epsilon_t, \qquad (3.15)$$

μ_t now containing spectral intercepts and a linear trend. Notice that (3.15) is identical to the corresponding second stage regression from a Dickey–Pantula sequence applied to $S(L)y_t$. In theory we would not reject $\mathsf{H}_{0,1}^1 : \pi_{1,1} = 0$.

In the context of (3.14), we are testing the null of second-order integration at the zero frequency and of first-order integration at the seasonal frequencies. As a consequence, the joint test for the exclusion of $\{\Delta_1 y_{j,t-1}\}_{j=1}^4$ will no longer have the same interpretation, but will have the same limiting null distribution, as when the notional upper bound on the order of integration is set common to all frequencies.

In Theorem 1 we give the asymptotic representations of the sequential test-statistics detailed above. We show that under $\mathsf{H}^r_{0,l}$, $l = 1, 2$ and $\cap^2_{i=1}\mathsf{H}^r_{0,2j+i}$, $j = 1, \ldots, S/2 - 1$, respectively, $r \in \{1, \ldots k\}$, the limiting distributions of the test-statistics, $t_{l,r}$, $l = 1, 2$, and of the $\{t_{2j+i,r}\}^2_{i=1}$ and $F_{2j+1\ 2j+2,r}$, $j = 1, \ldots, S/2 - 1$, are identical to those for the corresponding conventional seasonal unit root tests. Moreover, the $t_{j,d}$, $j = 1, \ldots, S$ and $F_{j,d}$, $j = 1, \ldots, S/2 - 1$, test-statistics diverge to either positive or negative infinity for $d > r$; that is, the tests are consistent under the alternative. The limiting distribution theory given in Theorem 1 applies whether we adopt a common notional upper bound on the maximum order of integration, or allow the upper bound to vary across frequencies, as in Example 1. In the latter case the notional upper bound must be at least as large as the true (unknown) order of integration at that frequency. In the former case it must be at least as large as the highest order of integration, taken across all frequencies.

Theorem 1. *In each case for $r \in \{1, \ldots, k\}$ and for deterministic component μ_t, denoting weak convergence by "$\Rightarrow$",*

(1) *Under $\mathsf{H}^r_{0,j}$, $j = 1, 2$, (a) $t_{j,r} \Rightarrow t^\mu_j$, and (b) $t_{j,d}$ diverges to negative infinity for $d > r$.*

(2) *Under $\cap^2_{i=1}\mathsf{H}^r_{0,2j+i}$, $j = 1, \ldots, S/2 - 1$, (a) $t_{2j+i,r} \Rightarrow t^\mu_{2j+i}$, $i = 1, 2$, and $F_{j,r} \Rightarrow \frac{1}{2}\left\{\sum^2_{i=1}(t^\mu_{2j+i})^2\right\}$, and (b) for $d > r$, $t_{2j+1,d}$ diverges to negative infinity, and/or $t_{2j+2,d}$ diverges to positive or negative infinity, while $F_{j,d}$ diverges to positive infinity.*

The notation t^μ_j refers to the asymptotic null distribution of the conventional ($r = k = 1$) seasonal unit root t-test, t_j, for $\mathsf{H}_{0,j} : \pi_j = 0$, when the auxiliary regression contains the deterministic component μ_t.

Remark 3. *The limiting distributions referred to in Theorem 1 are non-standard and are detailed in ST3. Finite-sample tabulations, relevant to $k = 1$, of these distributions are given for $S = 4$ in HEGY (Tables 1a and 1b, pp. 226–227), and ST1 (Tables 1a and 1b, p. 276), and for $S = 12$ in BM (Table A.1, pp. 325–26) and Taylor (1998, Tables I and II, pp. 356–57).*

The proof of Theorem 1 can be established from existing results in the literature. Consider first the sequence of zero-frequency tests, $t_{1,i}$, $i = k, \ldots, 1$. The sequence of HEGY regressions (that is, (3.5) for Step 1, (3.12) for Step 2, and so on) may be re-parametrized to (3.4); see Chan and Wei (1998, Appendix 1). Consequently, the proof for the zero frequency tests follows from the Appendix to Pantula (1989); see also Theorem 3.5.1 of Chan and Wei (1998). Non-stationary components at the seasonal frequencies are manifested in the lagged-dependent terms in (3.4). It is therefore important that the lagged-dependent variables are included in (3.4), in order to preserve the limiting null representations; see also, GLN (p. 420). Similar arguments apply to the sequences of Nyquist and harmonic seasonal frequency tests, with the limiting representations following along the same line of reasoning as in the Appendix to Pantula (1989). The limiting null representations given for the joint tests $F_{j,i}$, $j = 1, \ldots (S/2) - 1$, $i = 1, \ldots, k$, follow directly from the asymptotic orthogonality of the relevant regressors under the null hypothesis; see ST2 (Appendix). The consistency of the individual $t_{j,d}$, $j = 1, \ldots, S$, tests ensures the consistency of the $F_{j,d}$, $j = 1, \ldots, (S/2) - 1$, statistics under their respective alternatives.

Remark 4. *Joint frequency F-tests may be formed, as discussed in Section 3, and their asymptotic distribution under the nullity of the parameters under test will be given by the average of the null distributions of the corresponding squared t-statistics, again following directly from the asymptotic orthogonality of the relevant regressors under the null. Again the consistency of the individual t-tests guarantees the consistency of such joint tests.*

Remark 5. The limiting representations given in Theorem 1, remain asymptotically valid even if the disturbance process $\{\epsilon_t\}$ of (2.1) follows a stationary and invertible ARMA process with innovations satisfying the conditions in Pantula (1989). This result is conditional upon the sequence of auxiliary regressions being augmented with the lagged-dependent variables $\{\Delta_S^k y_{t-j}\}_{j=1}^{m_T}$, where m_T is a function of T, such that m_T^{-1} and $T^{-1}m_T^3$ both tend to zero as T tends to infinity (see Said and Dickey (1984), Pantula (1989)). The decomposition of $\phi_p(L)$ given in (2.2) is thus replaced by the approximation, $\phi_p(L) \approx \alpha(L)\beta_{m_T}(L)$, where $\beta_{m_T}(L)$ is of order m_T. In practice, however, the order of $\beta_{m_T}(L)$, and indeed of $\alpha(L)$, will be unknown and must be chosen by the practitioner. Moreover, the two choices are clearly inter-related. Suppose, for example, we fix $k = 2$ across all frequencies, but in fact the data follows the non-seasonal random walk, $\Delta y_t = \epsilon_t$. The first stage regression (3.5) imposes the (false) maintained hypothesis that the process is integrated of at least order one at all frequencies. Consequently, (3.5) does not admit a convergent $AR(\infty)$ representation. We therefore propose, following a referee's suggestion, to select the (lag truncation) order of $\phi_p(L)$ by fitting an AR model to y_t, using the data-based approach of Ng and Perron (1995). As in Ng and Perron (1995), we specify two parameters p_{max} and p_{min}, respectively maximum and minimum bounds on the lag truncation order. The maximum bound has to satisfy the conditions given above on m_T, while p_{min} is set equal to the notional upper bound on the order of $\alpha(L)$: e.g. if k is set common to all frequencies, then $p_{min} = kS$. Provided p_{min} is not less than the true order of $\alpha(L)$, the lag specification tests will have standard limiting distributions. Consequently, by imposing $m_T = \hat{p} - p_{min}$, $\hat{p}$ being the selected lag truncation order, on each regression in our testing sequence, Theorem 1 remains valid. One might also impose (zero) restrictions on some of the lagged difference terms in the test regressions (to obtain subset autoregressions) so that the adopted lag augmentation is parsimonious; see, inter alia, e.g. Hylleberg (1995) and Section 3 (pp. 318–319) of BM.

Recall that if A_T and B_T are two sequences of events on a probability space with limiting probabilities α and 1, respectively, then the limiting probability of $(A_T \cap B_T) = \alpha$. Consequently, and from Theorem 1, it follows that if we run our sequence of $t_{j,i}$, $i = k, \ldots, 1$, $j = 1, 2$ (where it is to be recalled that $j = 1$ indicates the zero frequency and $j = 2$ denotes the Nyquist frequency), test-statistics, each conducted at the α significance level, then

$$\lim_{T \to \infty} P_{H_{0,j}^r}\left(\text{rejecting } P_{H_{0,j}^d}\right) = \alpha, \qquad d = r$$
$$\leq \alpha, \qquad d < r$$
$$= 1, \qquad d > r. \qquad (3.16)$$

That is, using our sequential procedure, asymptotically, the likelihood that we would reject the hypothesis of exactly d unit roots at the zero (Nyquist) frequency is: (i) α, if there are exactly d unit roots at that frequency; (ii) no larger than α if there are more than d unit roots, and (iii) unity, if there are less than d unit roots. Identical limiting probabilities obtain for the sequences of joint F-tests, $F_{j,d}$, $j = 1, \ldots, (S/2) - 1$, conducting each test at the α significance level, when the true number of unit roots at jth harmonic frequency is r. Consequently, our sequential approach yields a consistent level-α testing procedure at each frequency. Size levels may be controlled across all frequencies by using the joint frequency tests outlined above (Section 3 and Remark 2). Furthermore, since all individual frequency tests are asymptotically independent under the null (see ST2), one may choose an exact significance level in the individual frequency tests, such that the overall size of the procedure (across all frequencies) is controlled, asymptotically.

Furthermore, if we let $\hat{r}_{j,\alpha}$, $j = 1, 2$, be an estimator of $r_{j,0}$, the true (but unknown) number of unit roots present at the zero ($j = 1$) and Nyquist ($j = 2$) frequencies, which takes the value d if $t_{j,d} \geq t_j^*(\alpha, T)$ and $t_{j,r} < t_j^*(\alpha, T)$, $r = d + 1, \ldots, k$ and takes the value zero when $t_{j,r} < t_j^*(\alpha, T)$, $r = 1, \ldots, k$, then from Theorem 3.1, it follows that, for $r_{j,0} > 0$,

$$
\lim_{T \to \infty} P_{\mathsf{H}_{0,j}^{r_{j,0}}} \left(\hat{r}_{j,\alpha} = r \right) = 1 - \alpha, \qquad r = r_{j,0}
$$
$$
\leq \alpha, \qquad r < r_{j,0}
$$
$$
= 0, \qquad r > r_{j,0}, \tag{3.17}
$$

$j = 1, 2$, and $P_{\mathsf{H}_{0,j}^0} \left[\hat{r}_{j,\alpha} = 0 \right]$, $j = 1, 2$, converges in probability to one. Identical asymptotic consistency results obtain for the seasonal harmonic frequency tests.

4. EMPIRICAL APPLICATIONS

In this section we apply our sequential method to two seasonally unadjusted macroeconomic time series: quarterly M4 in the UK, and monthly industrial production in the US. Both data-sets may be downloaded from `http://www.few.eur.nl/few/people/franses`.

4.1. Quarterly M4 in the UK

The first time series is quarterly M4 in the UK, observed for 1963.1–1988.4. This data-set has previously been analysed in Osborn (1990), where the source and an exact definition of the series is given. Osborn (1990) found, using the standard OCSB and HEGY methods, that the logged data required the Δ_1^2 transformation to remove two zero-frequency unit roots.

Denoting $\{y_t\}$ as the logged M4 time series, and setting $k = 2$, we first followed the approach detailed in Remark 5 and fitted an AR model to y_t, with $p_{\min} = 8$ and $p_{\max} = 16$ and using a 0.05 significance level in our lag specification tests. The lag order $\hat{p} = 10$ was deemed statistically acceptable with all standard mis-specification diagnostics easily passed. In what follows, unless otherwise stated, all regression diagnostic tests did not yield significant rejections. We therefore move on to consider the first-stage auxiliary regression

$$
\phi_{2,2}(L) \Delta_4^2 y_t = \mu_t + \sum_{j=1}^4 \pi_{j,2} \Delta_4 y_{j,t-1} + \epsilon_{t,1}, \tag{4.1}
$$

where the $\{y_{j,t}\}_{j=1}^4$ are as defined in (2.5), and μ_t is derived from special case (iv) of (2.1) and hence contains spectral dummies. Following from the specified AR model for y_t, $\phi_{2,2}(L)$ is a second-order lag polynomial. OLS estimation of (4.1) yielded the following test results, $t_{1,2} = -2.61$, $t_{2,2} = -7.00$, $F_{1,2} = 50.53$ and $F_{234,2} = 47.17$. Both lagged dependent variables were highly significant. Comparing these statistics with the relevant critical values in HEGY and GLN, all but the $t_{1,2}$ statistic is statistically significant at the 0.05 level. On the basis of this evidence we conclude that $\{\Delta_4 y_t\}$ contains a unit root at the zero frequency, but not at the seasonal frequencies. We thus move to the second-stage auxiliary regression,

$$
\phi_{2,1}(L) \Delta_4^2 y_t = \mu_t + \sum_{j=2}^4 \pi_{j,2} \Delta_4 y_{j,t-1} + \sum_{j=2}^4 \pi_{j,1} \Delta_1 y_{j,t-1} + \epsilon_{t,2}, \tag{4.2}
$$

where we note that μ_t still comprises spectral intercepts, and that $\phi_{2,1}(L)$ is of order two. OLS estimation of (4.2) yields $t_{2,1} = -1.35$, $F_{1,1} = 6.04$ and $F_{234,1} = 4.52$. Comparing these results with the tabulated critical values one finds that both F-type tests are highly significant, whilst the t-type test is insignificant at the 0.05 level. Within our framework of analysis, the logged M4 series therefore appears to require application of the $(1 - L)^2(1 + L)$ filter. Notice that this conclusion differs from that of Osborn (1990) who found no evidence for seasonal unit roots.

4.2. Monthly Industrial Production in the USA

As a second example, we consider the monthly industrial production index (1985 = 100) for the US, observed from May 1960 to July 1995; the data are taken from the OECD Macroeconomic Indicators database.

Denoting $\{y_t\}$ as the logged industrial production series, and again setting $k = 2$, we first fitted an AR model to y_t, with $p_{min} = 24$ and $p_{max} = 42$ and using a 0.05 significance level in our lag specification tests. The lag order $\hat{p} = 39$ was deemed statistically acceptable. We therefore considered the first-stage auxiliary regression

$$\phi_{15,2}(L)\Delta_{12}^2 y_t = \mu_t + \sum_{j=1}^{12} \pi_{j,2}\Delta_{12}y_{j,t-1} + \epsilon_{t,1}, \tag{4.3}$$

where: μ_t is derived from special case (iv) of (2.1) and thus comprises spectral dummies; $\phi_{15,2}(L)$ is of order 15, and the $\{y_{j,t}\}_{j=1}^{12}$ are as defined in pp. 307–308 of BM. To save space, we report only the zero- and biannual-frequency t-tests together with the joint-frequency F-tests of Taylor (1998). The most relevant test results from the OLS estimation of (4.3) are that $t_{1,2} = 2.32$, $t_{2,2} = -6.84$, $F_{2...12,2} = 46.10$ and $F_{1...12,2} = 46.94$. Retaining only the statistically significant lagged dependent variables (lags 1, 2, 12, 13 and 14) the corresponding results from (4.3) were: $t_{1,2} = -3.38$, $t_{2,2} = -10.70$, $F_{2...12,2} = 48.58$ and $F_{1...12,2} = 50.01$. All these statistics, with the exception of $t_{1,2}$ in the model with all lags included, imply the rejection of the corresponding null hypothesis at the 0.05 level, comparing with the relevant critical values in BM and Taylor (1998). We take the route implied by the $t_{1,2}$ statistic from the parsimonious lag model[†] and therefore consider the second-stage regression,

$$\phi_{15,1}(L)\Delta_{12}^2 y_t = \mu_t + \sum_{j=1}^{12} \pi_{j,2}\Delta_{12}y_{j,t-1} + \sum_{j=1}^{12} \pi_{j,1}y_{j,t-1} + \epsilon_{t,2}, \tag{4.4}$$

where μ_t now contains spectral intercepts and a time-trend. Again only the significant lagged variables $\Delta_{12}^2 y_{t-j}$, $j = 1, 2, 12, 13, 14$, were included. Notice, from (3.11), that (4.4) is equivalent to the auxiliary regression adopted in BM; that is (3.5) for $k = 1$ and $S = 12$. The key results from estimating (4.4) are that $t_{1,1} = -1.84$, $t_{2,1} = -3.20$, $F_{2...12,1} = 2.95$ and $F_{1...12,1} = 2.99$. The two joint F-type tests and the $t_{1,1}$ test statistic are all insignificant at the 0.05 level, while the $t_{2,1}$ test is significant at the 0.05 level, but not at the 0.01 level. Hence, and noting the issues surrounding multiple testing problems (see Section 3 and, for a detailed discussion, Taylor (1998)

[†]We also investigated the alternative route arising from acceptance of the hypothesis of second-order integration at the zero frequency. Conclusions on the seasonal frequencies were unchanged. Notice, therefore, that none of the test results from any of the regressions contradict the maintained hypothesis relevant to each model.

logged industrial production requires application of the Δ_{12}, or possibly the $\Delta_1 \Delta_{12}$, filter to remove unit roots at both the zero and seasonal frequencies, retaining a slight reservation regarding the single unit root at the bi-annual frequency. Moreover, this evidence suggests that application of the double $(1 + L + \cdots + L^{12})^2 \equiv S(L)^2$ filter to this variable would lead to non-invertibility, at all seasonal frequencies. One may therefore expect that the seasonally adjusted industrial production series will possess unit roots in its moving average representation.

5. CONCLUDING REMARKS

In this paper we have proposed a general method to examine the unit root properties of seasonal time series, allowing for the possibility of multiple unit roots at all or some of the zero and seasonal frequencies. The proposed sequential method is simple to compute, requiring no new tables of critical values. The sequential method has been shown to provide a consistent level-α testing procedure.

We have illustrated our approach using quarterly UK M4 and monthly US industrial production data. With these example series we have illustrated the relevance of our suggested procedure for examining, (i) the possibly inappropriate inferences drawn from existing methods, and (ii) the likely non-invertibility of seasonally adjusted data. Further research on these issues should reveal whether our current findings can be generalized across other macroeconomic data.

6. ACKNOWLEDGEMENTS

REFERENCES

Beaulieu, J. J. and J. A. Miron (1993). Seasonal unit roots in aggregate U.S. data, *Journal of Econometrics 55*, 305–328.

Bowerman, B. L., A. B. Koehler and D. J. Pack (1990). Forecasting time series with increasing seasonal variation, *Journal of Forecasting 9*, 419–436.

Clements, M. P. and D. F. Hendry (1997). An empirical study of seasonal unit roots in forecasting, *International Journal of Forecasting 13*, 341–355.

Chan, N. H. and C. Z. Wei (1988). Limiting distributions of least squares estimates of unstable autoregressive processes, *Annals of Statistics 16*, 367–401.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association 74*, 427–431.

Dickey, D. A., D. P. Hasza, and W. A. Fuller (1984). Testing for unit roots in seasonal time series, *Journal of the American Statistical Association 79*, 355–367.

Dickey, D. A. and S. G. Pantula (1987). Determining the order of differencing in autoregressive processes, *Journal of Business and Economic Statistics 5*, 455–461.

Engle, R. F., C. W. J. Granger, S. Hylleberg and H. S. Lee (1993). Seasonal cointegration: the Japanese consumption function, *Journal of Econometrics 55*, 275–298.

Franses, P. H. (1991). Seasonality, nonstationarity and the forecasting of monthly time series, *International Journal of Forecasting 7*, 199–208.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*, Second Edition, Wiley: New York.

Ghysels, E., H. S. Lee and J. Noh (1994). Testing for unit roots in seasonal time series: Some theoretical extensions and a Monte Carlo investigation, *Journal of Econometrics 62*, 415–442.

Haldrup, N. (1994). The asymptotics of single equation cointegration regression models with I(1) and I(2) variables, *Journal of Econometrics 63*, 153–181.

Hannan, E. J., R. D. Terrell and N. Tuckwell (1970). The seasonal adjustment of economic time series, *International Economic Review 11*, 24–52.

Hasza, D. P. and W. A. Fuller (1982). Testing for nonstationary parameter specifications in seasonal time series models. *Anals of Statistics 10*, 1209–1216.

Hylleberg, S. (1995). Tests for seasonal unit roots: general to specific or specific to general?, *Journal of Econometrics 69*, 5–25.

Hylleberg, S., R. F. Engle, C. W. J. Granger adn B. S. Yoo (1990). Seasonal integration and cointegration, *Journal of Econometrics 44*, 215–238.

Laroque, G. (1977). Analyse d'une methode de desaissonnalisation: le programme X-11 due Bureau of Census, version trimestrielle, *Annales de l'INSEE 28*, 105//127.

Maravall, A. (1995). Unobserved components in economic time series, in H. Pesaran and M. Wickens (eds.), *Handbook of Applied Econometrics*, Oxford: Basil Blackwell.

Ng, S. and P. Perron (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag, *Journal of the American Statistical Association 90*, 268–281.

Osborn, D. R. (199). A survey of seasonality in UK macroeconomic variables, *International Journal of Forecasting 6*, 327–336.

Osborn, D. R., A. P. L. Chui, J. P. Smith and C. R. Birchenhall (1988). Seasonality and the order of integration for consumption, *Oxford Bulletin of Economics and Statistics 50*, 361–377.

Pantula, S. G. (1989). Testing for unit roots in time series data, *Econometric Theory 5*, 256–271.

Said, S. E. and D. A. Dickey (1984). Tests for unit roots in autoregressive-moving average models of unknown order, *Biometrika 71*, 599–609.

Sen, D. L. (1985). Robustness of single unit root test statistics in presence of multiple unit roots, Ph. D. disseration, Department of Statistics, North Carolina State University.

Smith, R. J. and A. M. R. Taylor (1998). Additional critical values and asymptotic representations for seasonal unit root tests, *Journal of Econometrics 85*, 269–288.

Smith, R. J. and A. M. R. Taylor (1999a). Regression-based seasonal unit root tests, Working Paper 99-15, Department of Economics, University of Birmingham.

Smith, R. J. and A. M. R. Taylor (1999b). Likelihood ratio tests for seasonal unit roots, *Journal of Time Series Analysis 20*, 453–476.

Taylor, A. M. R. (1998). Testing for unit roots in monthly time series, *Journal of Time Series Analysis 19*, 349–368.